Multiple Linear Regression Techniques in Assessing the Effects of Truncation on the Gamma Distribution

Gannat Ramadan Mohamed El-Daian
Al-Azhar University for Girls, Egypt

Samuel R. Houston
University of Northern Colorado

Robert L. Heiny
University of Northern Colorado

This study examines, through Monte Carlo simulation, the effects of truncation on the gamma distribution (with one through four parameters). Specifically, three types of truncation (right, left, and double truncation) are considered. Computer facilities were used to generate 400 random samples from the gamma distribution with different parameter values for different sample sizes, shape, scale, location, power, and degree of truncation. Correlation and regression analysis demonstrated that the degree of left truncation has a significant correlation with the measures of central tendency for all distributions. The degree of right truncation had a significant relationship with the measures of deviation. On the other hand, the kind of truncation had a significant unique contribution for all models. However, the type of truncation had significant unique contributions for all models except two (scale and location models for gamma with three parameters).

The gamma distribution, or what may be called Pearson Type III of frequency curves, is one of the most important statistical distributions. It has been studied and investigated by many writers because of its application in different areas such as industrial engineering, physics, and quality control. For example, the gamma distribution can be considered as a description of duration variables such as the time taken for an instrument to be repaired, the time taken to get served at a store, etc.

If x is a continuous random variable having the gamma distribution, the generalized gamma density function is given by:

$$f(x; c, \beta, \lambda, \alpha) = \left(\beta\lambda^{-\alpha\beta}/\Gamma(\alpha)\right)(x-c)^{\alpha\beta-1}\exp\left(-((x-c)/\lambda)\beta\right) \quad [1]$$

where $0 \leq c < x < \infty$, $\beta > 0$, $\lambda > 0$, $\alpha > 0$, c is the location parameter, λ is the scale parameter, β is the power parameter, and α is the shape parameter.

The gamma distribution may have one, two, three, or four parameters. The probability density function for a gamma distribution with three parameters may be obtained from the generalized form (gamma with four parameters) by letting $\beta = 1.0$. The probability density function with two parameters can be obtained by using $c = 0.0$ and $\beta = 1.0$. The probability density function with one parameter can be obtained by letting $c = 0.0$ and $\beta = \lambda = 1.0$ in the generalized form.

In some situations, the complete range of the gamma distribution is not available to the researcher, in which case he or she works with a truncated distribution. The general form of the probability density function in this case becomes:

$$f(x|a < x < b) = \frac{f(x)}{\int_a^b f(x)dx} \quad [2]$$

where $a < x < b$ and where the values of a and b are dependent on the type of truncation and truncation degree. Therefore, if the range of x in (2) is $[a,\infty)$, the distribution can be called a left truncated gamma distribution. However, if the range of x is $(0,b]$, the distribution can be called a right truncated gamma distribution.

Many writers have discussed the subject of truncated distributions. Most of these studies are devoted to the exponential and normal distribution (Basu, 1964a; Yang & Sirvanci, 1977; Megahed, 1981; Depriest, 1983; Mittal, 1984; Mittal & Dahiya, 1987). Some writers have discussed the subject of the truncated gamma distribution. Most of these studies are devoted to the estimation of the parameters by the method of moments or by the method of maximum likelihood. Cohen (1950) used the method of moments while Des Raj (1953), Broeder (1955), and Chapman (1956) used the method of maximum likelihood.

In this paper, three kinds of truncation were considered: (a) degree of truncation is defined as size or degree of right, left, and double truncation ((tr), (tl) and (tl,tr)), (b) type of truncation is binary coded, where right truncation is (0,1), left truncation is (1,0) and double truncation is (1,1), and (c) interaction is coded tl*tr(0,1) and tr*tl(1,0).

Truncation can be one-sided (right or left) or two-sided. The three main types of truncation are right truncation, left truncation, or double truncation. In each case the fraction of the distribution of the population that falls outside the truncation point or points is called the degree of truncation. In this study two different approaches were used to analyze the effects of truncation on the gamma distribution. First, correlation analysis was used to determine the effect of truncation (type and degree) on the measures of central tendency and/or variation for the gamma distributions. Second, the multiple linear regression approach was used to generate and analyze models to isolate unique contributions of the different kinds of truncation.

Briefly, the major purpose in using multiple linear regression was to determine the actual impact of truncation (type, degree, etc.) on the gamma distribution of one, two, three and four parameters. The measures of central tendency are changed as a result of the type of truncation and/or its degree. The regression technique assisted in defining the theoretical consequences of truncation in the gamma distribution.

Research Limitations

This study was carried out using empirical data generated by a Monte Carlo simulation. The data were used to investigate the effects of different kinds of truncation on the gamma distribution. Computer facilities at the University of Northern Colorado were used to generate 400 random samples from the gamma distribution with different parameter values. The generated sample sizes were 10, 20, 30, 50, and 100. The shape parameter (α = 0.5, 1.5, 2.0, 3.0, 5.0 and 10.0), the scale parameter (λ = 1.0, 1.5, 2.0, 3.0, 4.0, and 5.0), the location parameter (c = 0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0, 3.0, and 5.0), and the power parameter (β = 1.0, 2.0, 3.0, 4.0, and 5.0) were considered. The truncation degrees were t times α, where t = 0.1, 0.2, 0.5, 1.0, 2.0, and 3.0.

Generation Of Gamma Variables

Since the gamma forms are easily obtained from raw data, they can approximate a wide variety of functional shapes. It could play a major role in digital simulation studies. Various investigators have been concerned with generating gamma variables. Phillips and Beightler (1972) presented a technique for generating random gamma varieties depending on two parameters:

$$f(x) = \left(\lambda^{-\alpha}/\Gamma(\alpha)\right)x^{\alpha-1}e^{-x/\lambda}, \quad 0 < x < \infty$$
$$f(x) = 0 \qquad \text{elsewhere} . \qquad [3]$$

They made a comparison between the composition technique and the rejection technique. With regard to statistical goodness-of-fit based on limited experiments, the first three methods were capable of generating random gamma variables closely approximating the desired gamma density for values of $\alpha > 1$. For lower values of the scale parameter λ, their method was better than the others. With regard to computer generation times, the rejection method is recommended for values of $\alpha < 2.5$. Naylor's (1971) method did quite well for values of $\alpha > 1$; although it was an easy method to program, it did require more computer running time.

Whittaker (1974) introduced a method of generating random variables from uniform variables with a gamma or beta distribution having a non-integer shape parameter. Cheng (1977) described the rejection method for generating gamma varieties with shape parameter α where $\alpha > 1$. The scheme used to generate gamma varieties with one parameter is described below.

1. Set $a = (2\alpha - 1)^{-1/2}$, $b = \alpha - \log 4$ and $c' = \alpha + \alpha^{-1}$.
2. Generate a pair of uniform random numbers, U_1 and U_2.
3. Set $v = a\log(U_1/(1-U_1))$, $x = \alpha\ e^{v}$.
4. If $b+c'v-x \geq \log(U_1^2U_2)$, accept x, otherwise go to Step 2.

This method was better than the previously published method (Phillips & Beightler, 1972; Whittaker, 1974) in terms of speed and program compactness.

Cheng and Feast (1979) introduced a simpler and faster method for generating a gamma variate by using the rejection technique. This method is suitable for all $\alpha > 1$ and it uses the ratio of uniform variable. It was noticed that some studies for generating gamma variable depend on specified conditions. Other methods need more time because the expected number of trials needed for each accepted variate could be more complicated than anticipated, and the complexity of the calculations required per trial are more extensive than anticipated. This is usually determined largely by the number of uniform random numbers needed and by the number of logarithmic or exponential function evaluations required.

Ripley's (1983) algorithm was developed on the basis of a recent study to generate gamma variables. It can be executed by using a small computer. However, it was found that the results were not suitable for most values of α and the generated variables did not follow the gamma distribution when $\alpha > 3$. The Ripley algorithms are called GKM1, GKM2, and GKM3. Algorithm GKM1 is described for $1 < \alpha < 4$. Algorithm GKM2 applies for $\alpha > 4$. More importantly, they are composite algorithm, called GKM3, which remains correct while covering all $\alpha > 1$.

Main Steps of Algorithm GKM1

1. Set $a = \alpha - 1$, $b = (\alpha - (6\alpha)^{-1}) / a$, $c' = 2 / a$ and $d = c'+2$.
2. Generate independent U(0,1) variates U_1 and U_2.
3. Let $w = bU_1/U_2$. If $(c'U_2 - d+w+w^{-1}) \leq 0$ go to 4.
4. If $c' \log U_2 - \log w+w^{-1} \geq 0$ go to step 1.
5. Deliver $x = aw$.

Main Steps of Algorithm GKM2

1. Let a, b, c', d as in GKMI; $f = \sqrt{\alpha}$.
2. Generate independent U (0,1) variable U_1 and U. Set $U_2 = U_1 + f^{-1}(1-1.86U)$. Repeat this step unless $0 < U_2 < 1$ (the constant f must be less than $(1+\sqrt{2}/e)$).

Steps 3, 4, and 5 are as in GKM1.

This method is complementary to GKM1 in that it is slower than GKMI for a near 1, but it rapidly becomes faster as a increases. Since GKMI and GKM2 differ only in step 2, it is easy to combine them by use of a switch. The composite algorithm is called GKM3.

Main Steps of Algorithm GKM3

1. This is exactly the same as GKM2 except that step 2 is replaced by 2'.

2'. Use step 2 of GKMI or GKM2 according to whether α is less than or greater than a prescribed value α_0. The suggested value for α_0 is 2.5; this ensures that the speed of variate generation is substantially the same for all $\alpha > 1$.

Three algorithms are needed for the three separate programs; therefore, it is preferred that the three algorithms be combined to make a modified algorithm which will be suitable for all values of α. In this way, a program is developed which can be generalized for generating gamma random variables with three or four parameters for any value of α, and value of λ, β, or c. The modified algorithm is introduced in the next section.

Research Design And General Procedures

The Faculty of Commerce Computer facilities at Al-Azhar University for Girls were used first to apply the Monte Carlo simulation. The Pseudo-Random Number Subroutine from the alphatronic Microcomputer System (BH/01824/e/k) was used next. Finally, the IBM 3081 G32 Computer System at the University of Northern Colorado was used.

Different random samples were used for generating gamma variables and different gamma distributions were generated with shape parameter (α = 0.5, 1.5, 2.0, 3.0, 5.0, and 10.0), scale parameter (λ = 1.0, 1.5, 2.0, 3.0, 4.0, and 5.0), location parameter (c = 1.0, 0.2, 0.3, 0.5, 1.0, 3.0, and 5.0), and power parameter (β = 1.0, 2.0, 3.0, 4.0, and 5.0). BASIC programs which were designed for generating gamma varieties were dependent on the following modified algorithm.

The researchers developed an algorithm through the mixture of GKMI, GKM2, and GKM3. The algorithm was developed to be suitable for all values of the parameters of the three parameter gamma distribution. Supposing that ($\alpha < 1$) is a shape parameter, the algorithm follows these steps:

1. Let $a = \alpha - 1$, $b = (\alpha - (6\alpha)^{-1}) / a$, $c' = 2 / a$, $d = c' + 2$ and $f = \sqrt{\alpha}$.
2. Generate independent U(0,1) variables U_1 and U_2.
3. If ($\alpha < 4$) go to step 5.
4. Set $U_2 = U_1 f^{-1}(1 - 1.86 U_0)$. Repeat this step unless $0 < U_2 < 1$.
5. Let $w = b U_1/U_2$. If $(c'U_2 - d+w+w^{-1}) \le 0$, go to step 7.
6. If $(c'\log U_2 - \log w + w - 1) \ge 0$ go to step 2.
7. Deliver $x = (aw)\lambda + c$.

Description of the Variables

The variables used for the analyses in this study are listed in Table 1.

Table 1 List of Variables

Parameters		
1	Shape Parameter	Y1
2	Scale Parameter	Y2
3	Location Parameter	Y3
4	Power Parameter	Y4
Measures of Central Tendency and Deviation		
5	Mean	$\bar{x}$
6	Variance	s^2
7	Skewness	Sk
8	Kurtosis	Ku
9	Mode	Mo
10	Median	Me
Sample Size		
11	Sample Size	n
The Kind of Truncation		
12	Degree lt truncation	tl.
13	Degree rt truncation	tr.
14	Left truncation	tl(1,0)
15	Right truncation	tr(0,1)
16	Interaction 1	tl*tr(0,1)
17	Interaction 2	tr*tl(1,0)
The Kind of Distribution		
18	Gamma one parameter	k1
19	Gamma two parameters	k2
20	Gamma three parameters	k3
21	Gamma four parameters	k4

Table 2 Relationships Between Central Tendency and Deviation Measures and Kind of Truncation for Gamma Distributions (N=250)

Measure	tl	tr	tl(1,0)	tr(0,1)	tl*tr(0,1)	tr*tl(1,0)
Mean						
One	.977*	-.011	.238*	-.347*	-.023	-.027
	(.0001)	(.857)	(.0001)	(.0001)	(.718)	(.677)
Two	.727*	-.439*	.199*	-.524*	-.122	-.297*
	(.0001)	(.0001)	(.002)	(.0001)	(.059)	(.0001)
Three	.139*	.306*	-.114	-.194*	.113	.089
	(.028)	(.0001)	(.072)	(.002)	(.075)	(.089)
Four	.403*	.314*	.078	-.269*	-.059	.171*
	(.0001)	(.0001)	(.219)	(.0001)	(.352)	(.007)
Mode						
One	.980*	-.030	.239*	-.345*	-.030	-.030
	(.0001)	(.641)	(.0001)	(.0001)	(.636)	(.539)
Two	.756*	-.423*	.223*	-.503*	-.103	-.285*
	(.0001)	(.0001)	(.0003)	(.0001)	(.109)	(.0001)
Three	.185*	.309*	-.062	-.199*	.150*	.123
	(.003)	(.0001)	(.329)	(.002)	(.017)	(.052)
Four	.412*	.302*	.081	-.278*	-.062	.159*
	(.000)	(.0001)	(.202)	(.0001)	(.325)	(.012)
Median						
One	.977*	-.011	.857*	-.347*	-.022	-.027
	(.0001)	(.858)	(.0001)	(.0001)	(.718)	(.676)
Two	.728*	-.438*	.200*	-.524*	.122	-.297*
	(.0001)	(.0001)	(.002)	(.0001)	(.054)	(.0001)
Three	.140*	.308*	-.114	-.192*	.114	.090
	(.027)	(.0001)	(.073)	(.002)	(.072)	(.155)
Four	.403*	.314*	.078	.269*	-.059	.171*
	(.0001)	(.0001)	(.221)	(.0001)	(.352)	(.007)
Variance						
One	.492*	.533*	.111	-.218*	.174*	.319*
	(.0001)	(.0001)	(.081)	(.0005)	(.006)	(.0001)
Two	.389*	-.419*	-.009	-.475*	-.189*	-.294*
	(.0001)	(.0001)	(.892)	(.0001)	(.003)	(.0001)
Three	-.136*	.145*	-.297	-.142*	-.120*	-.145*
	(.0001)	(.0001)	(.892)	(.0001)	(.003)	(.0001)
Four	.020	.193*	-.028	-.058	-.030	.106
	(.752)	(.002)	(.660)	(.363)	(.638)	(.095)
Skewness						
One	-.139*	-.222*	-.063	.009	-.022	--.127*
	(.028)	(.0004)	(.323)	(.891)	(.734)	(.044)
Two	.045	-.183*	.115	-.167*	.146*	.013
	(.476)	(.004)	(.071)	(.008)	(.021)	(.636)
Three	-.077	-.537*	.301*	-.347*	-.135*	-.202*
	(.228)	(.0001)	(.0001)	(.0001)	(.034)	(.001)
Four	.245*	.081	.104	-.158*	.344*	.186*
	(.0001)	(.204)	(.102)	(.012)	(.0001)	(.003)
Kurtosis						
One	-.138	-.200*	-.202*	.148*	-.038	.151*
	(.300)	(.002)	(.001)	(.020)	(.550)	(.017)
Two	-.015	.139*	.144*	.160*	.319*	.222*
	(.816)	(.028)	(.023)	(.011)	(.0001)	(.0004)
Three	-.054	.229*	.347*	.004	-.064	-.077
	(.397)	(.0003)	(.0001)	(.947)	(.315)	(.228)
Four	.227*	.213*	.052	-.041	.343*	.205*
	(.0003)	(.0007)	(.410)	(.524)	(.0001)	(.001)

Note. One, Two, Three, and Four are gamma distributions with one through four parameters; top numbers refer to Pearson correlation coefficients; numbers in parentheses refer to p-values; asterisks indicate a significant relationship ($p < .05$).

Correlation Analysis

Table 2 shows the correlation coefficients between the central tendency measures and the kind of truncation for the gamma distribution with one through four parameters. This table indicates that the degree of left truncation had the highest correlation with the measures of central tendency of the gamma distribution with one through four parameters.

The degree of right truncation had negative correlations with the measures of central tendency of the gamma distribution with one and two parameters. The degree of right truncation has a significant relationship with the measures of central tendency for gamma distributions except the gamma with one parameter.

Table 2 indicates the relationship between the type of truncation and the measures of central tendency. The relationships between the left truncation and the measures of central tendency were positive and significant for the gamma with one and two parameters. The table also indicates that right truncation had negative and significant correlations with measures of central tendency for gamma distributions with one through four parameters. Further, there was no evidence to support that the interaction tl*tr(0,1) had a significant correlation with the measures of central tendency for the gamma distributions except the mode of the three parameter gamma distribution. However, the interaction tr*tl(1,0) had significant correlations with measures of central tendency for the gamma with one and four parameters.

Also, Table 2 indicates that all types of truncation except the left truncation had a significant correlation with the variance of the gamma distributions with one through three parameters. The degree of left truncation has a significant correlation with the skewness of the gamma with one and four parameters, but there was no evidence to support that there were relationships with the skewness of the gamma with two and three parameters. The degree of right truncation had a significant correlation with the skewness of the gamma with one through three parameters.

Left truncation was not significantly correlated with the skewness of the gamma with one, two, and four parameters; however, it has a significant correlation with the skewness of the gamma with three parameters. The right truncation has a significant correlation with skewness in the different cases, except for the gamma with one parameter.

Degree of left truncation has significant correlations only with the kurtosis of the gamma with four parameters; however, the degree of right truncation has significant correlation with kurtosis of the gamma with one through four parameters. Left truncation has significant correlations with the kurtosis of the gamma with one and three parameters; however, right truncation has significant correlations with the gamma with one and two parameters. The interactions had significant correlation with the gamma with one and four parameters.

Table 3 presents the relationships between measures of central tendency, dispersion, distributional shape measures, and the kind of truncation. This table is based on all samples which are used for one through four parameter distributions, and is a summary of these relationships. It indicates that the relationships between the measures of central tendency and both the degree and type of truncation were significant.

There was no evidence that the degree of truncation was significantly related to the variance but the type of truncation was significantly related to the variance. The degree and the type of right truncation had a significant correlation with the measures of distributional shape, while there was no evidence to support that left truncation was significantly related to the measures of distributional shape.

Table 3 Relationships Between Distribution Measures and Kind of Truncation (N=1000)

Measure	tl	tr	tl(1,0)	tr(0,1)	tl*tr(0,1)	tr*tl(1,0)
Mean	.655	.081	.066	-.331	.032	-.009
	(.0001)	(.010)	(.036)	(.0001)	(.306)	(.774)
Mode	.697	.076	.105	-.329	.039	.001
	(.0001)	(.016)	(.0009)	(.0001)	(.217)	(.968)
Median	.655	.082	.067	-.330	.033	-.009
	(.0001)	(.010)	(.035)	(.0001)	(.301)	(.788)
Variance	.049	.034	-.190	-.183	-.064	-.111
	(.119)	(.277)	(.0001)	(.0001)	(.045)	(.0004)
Skewness	.005	-.267	.060	-.167	-.010	-.118
	(.870)	(.0001)	(.057)	(.0001)	(.762)	(.0002)
Kurtosis	-.044	-.135	.036	.072	.044	.071
	(.162)	(.0001)	(.254)	(.022)	(.153)	(.026)

Note. Top numbers refer to Pearson correlation coefficients; numbers in parentheses are p-values.

Table 4 Relationships Between Kind of Truncation and Number of Parameters as Coded by Binary Vectors (N=1000)

	Number of Parameters of Gamma Distribution			
Kind of Truncation	One	Two	Three	Four
tl	.088	.036	-.042	-.082
	(.005)	(.256)	(.181)	(.010)
tr	-.019	-.149	.088	.080
	(.547)	(.0001)	(.005)	(.011)
tl(1,0)	.077	-.030	-.169	.124
	(.014)	(.336)	(.0001)	(.0001)
tr(0,1)	.004	-.095	.047	.045
	(.912)	(.003)	(.140)	(.159)
ti*tr(0,1)	-.061	-.059	.152	-.033
	(.055)	(.060)	(.0001)	(.300)
tr*tl(1,0)	-.011	.094	-.026	.131
	(.731)	(.003)	(.415)	(.0001)

Note. Top numbers are Pearson correlation coefficients; numbers in parentheses are p-values.

Table 4 identifies the relationships between the kinds of truncation and the distributions. This table indicates that the gamma distribution with one parameter had significant relationships with the degree and type of left truncation. The gamma distribution with two parameters had significant relationships with the degree and type of right truncation and the interaction tr*tl(1,0). The gamma distribution with three parameters had significant relationships with the degree of right truncation, left truncation and the interaction (both left and right), left truncation and the interaction tr*tl(1,0).

Briefly, the relationships between truncation and the measures of gamma distributions were dependent on the kind of distribution, the type of truncation, and the distribution measures; whereas the relationships among the measures of central of tendency and the types or degrees of truncation were all significant.

Multiple Regression Analysis

Multiple regression analyses were performed to examine the relationships between each parameter of the gamma distribution (dependent variable) and the set of descriptive characteristics (independent variables) of the gamma distributions. Determination of which variable serves as criterion variable and which set of variables are predictor variables was dependent on the model that was considered.

In general, the set of descriptive characteristics of the gamma distributions includes the variables,

1. Measures of central tendency (mean, mode, and median) to be represented by X_1, X_5, and X_6, respectively.
2. Measures of deviation (variance, skewness, and kurtosis) represented by X_2, X_3, and X_4, respectively.
3. Sample size represented by X_7.
4. Degree of truncation (tr) and (tl), represented as X_8 and X_9, respectively.
5. Type of truncation tl(1,0) and tr(0,1), represented as X_{10} and X_{11}, respectively.
6. Interaction tl*tr(0,1) and tr*tl(1,0), represented as X_{12} and X_{13}, respectively.
7. Kind of gamma distribution with one through four parameters represented as X_{14}, X_{15}, X_{16}, and X_{17}, respectively. Each variable was binary coded.

The full model considers all seventeen variables (X_1, X_2, ... X_{17}) as predictors for the dependent variables (one of the four gamma distribution parameters).

Sample size should be related to the number of variables and should increase as the number of variables increases (Barcikowski & Stevens, 1975). One informal guide for a lower limit is that there should be 10 subjects for each variable. To insure sufficient sample size for a small set of variables, Thorndike (1978) offered the following rule to determine the sample size: $N \geq 10(p+c)+50$, where p is the number of independent variables (predictors) and c is the number of dependent variables (criteria). In this study, Thorndike's sample size requirement was satisfied in all regression models which were analyzed.

A hierarchical chart of reduced models was constructed (see Table 5). Each model was reduced from the full model and significant drops in R^2 indicated significant contributions to the dependent variable, Y, by omitted predictors are identified by asterisks using the .05 level of significance (see Table 6).

Also, comparisons between R^2 for the different models are summarized in Table 6 for the gamma with one through four parameters and the general model (which includes all the data of the gamma distribution with one through four parameters) for the shape of the parameter model. This table indicates that a sequential decrease in R^2 occurred from gamma with one through gamma with four parameters for the shape models (.952 to .733). R-squares for full models depended on the kind of distribution and N. Furthermore, the R^2 value in the general case for the four parameter model dropped to .733 (N=1000) from .868 (N=250). Thus, it can be seen that both the sample size and number of parameters were factors which influence the R^2 value.

Table 5 List of Regression Analysis Reduced Models

#	The Reduced Model
1	Full model - measures of central tendency (MCT)
2	Full model - measures of deviation (M.Dev)
3	Full model - sample size (n)
4	Full model - kinds of truncation
5	Full model - mean
6	Full model - mode (mo)
7	Full model - median (me)
9	Full model - distributional shape measure (ds)
10	Full model - skewness (sk)
11	Full model - kurtosis (ku)
12	Full model - (tr,tl)
13	Full model - (tl(1,0), tr(0,1)
14	Full model - interactions (int)
15	Full model - tl.
16	Full model - tr
17	Full model - tl (1,0)
15	Full model - tr (0,1)
19	Full model - kinds of distribution (for the general model)

Table 6 R-Square Values for Full and Reduced Models in Regression Analysis for Gamma Distribution Where Y is the Shape Parameter

	Number of Parameters of Gamma Distribution				
Model	One	Two	Three	Four	General
FM	.952	.948	.886	.868	.733
FM-MCT	.948*	.511*	.864*	.576*	.580*
-Mean	.952	.948	.885	.868	.732
-Mo	.952	.948	.879*	.865*	.732
-Me	.952	.948	.885	.868	.732
FM-Dev.	.946*	.810*	.865*	.860*	.714*
-Var.	.952	.812*	.873*	.861*	.717*
FM-Ms.	.947*	.942*	.884	.867	.731
-Sk	.949*	.942*	.885	.867	.733
-Ku	.948*	.943*	.884*	.868	.733
FM-n	.952	.948	.882	.867	.732
FM-Truncation	.897*	.498*	.809*	.626*	.519*
-degrees	.947*	.645*	.857*	.693*	.613*
-tl.	.948*	.663*	.886	.724*	.667*
-tr.	.951*	.948	.858*	.862*	.705*
-types	.947*	.931*	.872*	.854*	.722*
-tl(1,0)	.951*	.948	.881*	.867	.722*
-tr(0,1)	.949*	.931*	.879*	.855*	.732
-interaction	.924*	.848	.885	.818*	.709*

Note: Asterisk indicates that the unique contribution for these variables was significant at .05 level.

Table 7 Significant Unique Contributions for each Kind of Truncation in Regression Analysis Models Kind of Truncation

Y	tl	tr	tl(1,0)	tr(0,1)	Interaction
Shape Parameters					
One	*	*	*	*	*
Two	*			*	*
Three		*	*	*	
Four	*	*		*	*
General	*	*	*		*
Scale Parameters					
Two	*	*	*	*	*
Three		*			
Four	*		*	*	*
General	*	*			
Location Parameters					
Three		*			
Four	*			*	*
General	*	*		*	*
Power Parameters					
Four			*	*	
General	*				

Note. Asterisk indicates that the unique contribution for these variables was significant at the .05 level.

Moreover, Table 6 indicates that the measures of central tendency, deviation, and the kinds of truncation (and especially the types of truncation) had significant unique contributions. Most kinds of truncation had significant unique contributions for all models. The variance had a significant, unique contribution for all distributions except for the gamma with one parameter. Moreover, skewness and kurtosis had a significant unique contribution for the gamma with one and two parameters.

Briefly, Table 7 shows the significant unique contributions of each kind of truncation in each model. It indicates that the degree of left truncation had a significant unique contribution with all models except the models of the gamma with three parameters and the model of the power parameter in the gamma with four parameters. The degree of right truncation had significant unique contributions for all models except for the shape model for the gamma with two parameters, for location model for gamma with four parameters, and for the power models.

Left truncation had significant unique contributions with the shape models, except the general model, in the scale model for the gamma with two and three parameters, and in the power model for the gamma with four parameters. Right truncation had significant unique contributions with the shape models, except the general model, with the scale model for the gamma with two and four parameters model, with the location model for the gamma with four parameters, with the general gamma model, and with the power model for the gamma with four parameters.

The last kind of truncation (interaction) had significant unique contributions with the shape model for the gamma with two and four parameters, with the scale model for the gamma with two and four parameters, and with the location model for the gamma with four parameters.

Moreover, the types of truncation were considered as elements in the best proper subset for the different models. According to the backward elimination procedure for a dependent variable, Y, for each model, the best proper subset of the predictors for each model is given in Table 8.

In general, Table 8 indicates that it is possible that the best proper subset (using the backward elimination procedure) for the shape parameter included some type of truncation, for example, tl, tr, tl(1,0) and tl*tr(0,1) for the general model. For the shape parameter model of the gamma with four parameters, these same types of truncation patterns were included along with tr*tl(1,0). The best proper subset of independent variables for the scale parameter included tl, tr, and tl*tr(0,1) for the general model. For the shape parameter model of the gamma with four parameters, truncation variables including tl, tl(1,0), tr(0,1), and tl*tr(0,1) were identified as important. The best proper subset for the location parameter included tl, tr, tr(0,1), and tl*tr(0,1) for the general model. For the shape parameter model of the gamma with four parameters, it was found that tl, tr, tr(0,1), tl*tr(0,1), and tr*tl(1,0) were significant.

General Concluding Remarks

This study dealt with the effects of truncation on the family of gamma distributions. The goal was to determine whether the kind of truncation [tl, tr, tl(1,0), tr(0,1), tl*tr(0,1), and tr*tl(1,0)] had an influence on the

Table 8 The Best Proper Subset for Each Model in the Different Cases of Gamma Distributions

Model	General Model				Gamma with Four Parameters			
Subset/Y	Shape	Scale	Location	Power	Shape	Scale	Location	Power
1	mean	mean	Var	Var	Var	Var	mean	Sk
2	Var	Var	Sk	Sk	Mo	Mo	tl	Mo
3	tl	Sk	tl	Ku	tl	tl	tr	Me
4	tr	Ku	tr	Me	tr	tl(1,0)	tr(0,1)	
5	tl(1,0)	Me	tr(0,1)	tl	tr(0,1)	tr(0,1)	tl*tr(0,1)	
6	tl*tr(0,1)	tl	tl*tr(0,1)	tl*tr(0,1)	tl*tr(0,1)	tr*tl(1,0)	tr*tl(1,0)	
7	k1	tr	k1	k1	tr*tl(1,0)			
8	k2	tl*tr(0,1)		k2	k2			
9	k3	k1	k3	k3				
10		k2						
11		k3						
R^2	.729	.597	.875	.778	.865	.722	.845	.301

Model	Gamma with Three Parameters			Gamma with Two Parameters		Gamma with One Parameter
Subset/Y	Shape	Scale	Location	Shape	Scale	Shape
1	mean	var	mean	var	mean	mean
2	var	tl	var	sk	var	sk
3	mode	tl(1,0)	tl(1,0)	ku	ku	ku
4	n			Me	Me	tr
5	tr			tl	tl	tr(1,0)
6	tl(1,0)			tr	tr	tl*tr(0,1)
7	tr(0,1)			tr(0,1)	tr(0,1)	tr*tl(1,0)
8				tl*tr(0,1)	tl(1,0)	
R^2	.882	.589	.500	.948	.696	.948

parameters, central tendency, dispersion, and distributional shape measures of the gamma distributions.

From these analyses, it can be concluded that the truncation significantly affected the gamma distribution (its measures and its parameters). The effect was dependent on the kind of truncation (type and degree), the type of distribution, and the values of parameters.

The numerical analyses in this research presented abstract concepts about irregular relationships, but also introduced some details about the relationships between each kind of truncation and the parameters, and each kind of truncation and the most characteristic measures of gamma distributions. For example, the degree of left truncation had a significant correlation with the measures of central tendency for all kinds of distributions. On the other hand, the degree of right truncation had a significant relationship with the measures of central tendency and deviation for most kinds of gamma distributions.

In terms of multiple regression analysis, for all models, the kind of truncation had significant unique contributions. The degree of truncation made significant, unique contributions to all models, but the types of truncation had a significant unique contribution for all models except two (scale and location models for the gamma with three parameters). However, the unique contribution for each kind of truncation was dependent on the kind of truncation, the distribution, and the measure being considered.

The overall findings of this study were generally supportive of the findings of the reviewed research. Although this study was limited to the effects of truncation on the gamma distribution, it can be concluded that the findings apply also for Erlang, exponential, and chi-square distributions.

Suggestion for Further Studies

The focus of this study emphasizes the effects of truncation on the gamma distribution. Based upon the findings and conclusions described above, these recommendations are made for further investigation:

1. Results were dependent on specific values of the parameters. Therefore, for more generalizable results, different values of the parameters may be utilized in subsequent research.
2. Exploration should be continued in an effort to study the effect of truncation on other distributions such as beta, lognormal, Weibull, etc. Additional study could determine if the research findings for

different distributions are comparable for similar types of truncation. This could provide the researcher with information about the effects of truncation on different kinds of distributions. Also, research should be extended to include the study of effects of truncation on mixture distributions.

3. Moreover, the effects of truncation in both estimation and hypothesis testing when using the transformations mentioned by Mohamed (1981) should be examined.
4. The methods used in this study should be extended to examine the effects of inner truncation and partial truncation on the gamma distribution.
5. Other studies should be designed using other techniques such as canonical correlation or factor analysis to determine whether rotated and/or unrotated factor solutions are affected by the type and/or degree of truncation.

References

Barcikowski, R. S. & Stevens, J. P. (1975). A Monte-Carlo study of the stability of canonical correlation, canonical weights, and canonical variate-variable correlation. *Multivariate Behavioral Research, 10*, 353-364.

Basu, A. B. (1964a). Effect of truncation on a test for the scale parameter of the exponential distribution. *Annals of Mathmatical Statistics, 35*, 209.

Broeder, G. G. (1955). On parameter estimation for truncated pearson type III distribution. *Annals of Mathematical Statistics, 26*, 659-663.

Chapman, D. G. (1956). Estimating the parameters of truncated gamma distribution. *Annals of Mathematical Statistics, 27*, 498-506.

Cheng, R. C. H. (1977). The generation of gamma variables with non-integral shape parameter. *Applied Statistics, 26*, 71-75.

Cheng, R. C. H. & Feast, G. M. (1979). Some simple gamma variable generators. *Applied Statistics, 28*, 290-295.

Cohen, A. C. (1950). Estimating parameters of Pearson type III populations from truncated samples. *Journal of the American Statistical Association, 45*, 411-423.

Depriest, D. J. (1983). Using the singly truncated normal distribution to analyze satellite data. *Communication in Statistical Theory and Methods, 12*, 263-272.

Des Raj (1953). Estimation of the parameters of type III populations from truncated samples. *Journal of the American Statistical Association, 48*, 336-349.

Megahed, Abdel-Razek (1981). *A Monte-Carlo investigation of the normality assumption with truncated normal population.* Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.

Mittal, M. M. (1984). *Estimating the parameters of truncated distributions.* Unpublished doctoral dissertation, Old Dominion University, Norfolk, VA.

Mittal, M. M. & Dahiya, R. C. (1987). Estimating the parameters of a doubly truncated normal distribution. *Communication in Statistics/ Simulations, 16*, 141-149.

Mohamed, G. R. (1981). *Statistical implications of using a Gaussianting transformation.* Unpublished master's thesis. Al-Azhar University for Girls, Egypt.

Naylor, T. (1971). *Computer simulation experiments with models of economic systems.* New York, Wiley.

Phillips, O. T. & Beightler, C. S. (1972). Procedures for generating gamma varieties with non-integer parameter sets. *Journal of Statistical Computer Simulation, 1*, 197-208.

Ripley, B. D. (1983). Computer generation of random variable: A tutorial. *International Statistical Review, 51*, 301-319.

Thorndike, R. M. (1978). *Correlation procedures for research.* New York, Wiley.

Whittaker, J. (1974). Generating gamma and beta random variables with non-integral shape parameters. *Applied Statistics, 23*, 210.

Yang, G. & Sirvanci, M. (1977). Estimation of a time-truncated exponential parameter used in life testing. *Journal of the American Statistical Association, 72*, 444-447.